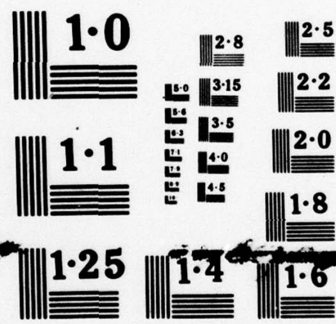
AD-A067 661 TEXAS UNIV AT AUSTIN DEPT OF ELECTRICAL ENGINEERING F/G 12/1
ESTIMATOR PERFORMANCE FOR A CLASS OF NONLINEAR ESTIMATION PROBL--ETC(U)
MAR 79 C LIU, S I MARCUS F49620-77-C-0101
UNCLASSIFIED AFOSR-TR-79-0454 NL

1 OF 1
ADA
067661

END
DATE
FILMED
6-79
DDC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

AFOSR-TR- 79-0454

LEVEL II

AD A067661

DDC FILE COPY

DDC RECEIVED APR 13 1979 C

Estimator Performance for a Class of Nonlinear Estimation Problems

Chang-Huan Liu and Steven I. Marcus**

ABSTRACT

The state estimation problem for a certain class of nonlinear stochastic systems with white Gaussian plant and observation noise is considered. The optimal (minimum variance) estimators for these systems are recursive and finite dimensional. A particular nonlinear system which contains a polynomial nonlinearity is presented. Both optimal and suboptimal estimators and an estimation lower bound for such a system are derived. The performance of the optimal and suboptimal estimators and the lower bound are compared both analytically and by computer simulation.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

*This research was supported in part by the Joint Services Electronics Program under Contract F49620-77-C-0101, in part by the National Science Foundation under grant ENG 76-11106, and in part by the Air Force Office of Scientific Research (AFSC) under grant AFOSR 79-0025.

**The authors are with the Department of Electrical Engineering, The University of Texas at Austin, Austin, Texas 78712.

Approved for public release; distribution unlimited.

Submitted to the IEEE Transactions on Automatic Control.

79 04 12 01

White Section
Buff Section
DISTRIBUTION/AVAILABILITY CODES
and/or SPECIAL

I. Introduction

Optimal estimators have been derived for very general classes of nonlinear systems [1],[2]. However, the optimal estimator requires, in general, an infinite dimensional computation to generate the conditional mean of the system state given the past observations. This computation involves either the solution of a stochastic partial differential equation for the conditional density or an infinite dimensional system of coupled ordinary stochastic differential equations for the conditional moments. For linear stochastic systems with linear observations and white Gaussian plant and observation noises, it is known that the optimal conditional mean estimator consists of a finite dimensional linear system (the Kalman-Bucy filter [3]). Many types of finite dimensional suboptimal estimators for nonlinear systems have been proposed [4], and many numerical experiments have been performed to compare the various suboptimal estimators. In addition, upper and lower bounds on the performance (error covariance) of optimal and suboptimal estimators have been derived for certain classes of nonlinear systems [5]-[10]. However, throughout most of the previous research, the question of how good the performance of a suboptimal estimator (or a lower bound) is, as compared to the performance of the optimal estimator (not as compared to that of other suboptimal estimators), has remained unexamined.

Recently, Marcus, Willsky, and Lo [11]-[15] have derived finite dimensional optimal nonlinear estimators for certain classes of nonlinear systems with white Gaussian plant and observation noise. The existence of such optimal estimators allows us to investigate how both optimal and suboptimal estimators process information, and to examine the differences between them. The classes of systems for which finite dimensional estimators are implementable are described by certain polynomial nonlinearities or by certain classes of Volterra series expansions. One class of models considered in [11]-[12] is described by the Ito equations

$$dx(t) = F(t)x(t)dt + G(t)dw(t) \tag{1}$$

$$\dot{y}(t) = (A_0 + \sum_{i=1}^{N} x_i(t)A_i)y(t) \tag{2}$$

$$dz(t) = H(t)x(t)dt + dv(t) \tag{3}$$

where $x(t)$ is an n-vector, $y(t)$ is a k-vector or k×k matrix, $\{A_i\}$ are k×k matrices, w and v are independent Brownian motion (Wiener) processes, and $x(0)$ is Gaussian. The optimal estimate, with respect to a wide variety of criteria, of $x(t)$ given the observations $z^t \triangleq \{z(s), 0 \le s \le t\}$, is the conditional mean $\hat{x}(t|t)$ (also denoted by $E^t[x(t)]$ or $E[x(t)|z^t]$) [4]. It is well known [3],[4] that the computation of $\hat{x}(t|t)$ can be performed by the finite dimensional (linear) Kalman-Bucy filter. However, the computation of $\hat{y}(t|t)$ requires in general an infinite dimensional system of stochastic differential equations. In [12], Marcus and Willsky have proved that $\hat{y}(t|t)$ is computable with a recursive finite dimensional estimator, if the ideal generated by A_0 in the Lie algebra $\{A_0, A_1, \ldots, A_N\}_{LA}$ is nilpotent. Using Volterra series expansions [16]-[17], Marcus and Willsky have also proved a similar result for systems described by (1),(3), and in which $y(t)$ is given by a certain type of Volterra series expansion or by polynomial type feed-forward nonlinearities [11]-[12], as illustrated by the following system:

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} -\alpha & 0 \\ 0 & -\beta \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} dt + q^{\frac{1}{2}} \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix} \triangleq F\,x(t)dt + q^{\frac{1}{2}}dw(t) \tag{4}$$

$$dy(t) = (-\gamma\; y(t) + x_1(t)x_2(t))dt \tag{5}$$

$$\begin{bmatrix} dz_1(t) \\ dz_2(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} dt + r^{\frac{1}{2}} \begin{bmatrix} dv_1(t) \\ dv_2(t) \end{bmatrix} \tag{6}$$

where α, β, $\gamma > 0$, w_1, w_2, v_1, and v_2 are independent, zero mean, unit variance Wiener processes, and $x_1(0)$, $x_2(0)$, and $y(0)$ are independent Gaussian random variables which are also independent of the noise processes. Here (5) involves a polynomial feed-forward nonlinearity.

In section II we construct, for the system (4)-(6), the optimal estimator and three suboptimal estimators: the extended Kalman filter (EKF), the constant gain extended Kalman filter (CGEKF), and the best linear estimator (BLE, also known as the linear minimum-variance estimator). Section III is concerned with the derivations of error covariance propagation equations for both optimal and suboptimal estimators and a lower bound based on that of Bobrovsky and Zakai [10] for the system (4)-(6). The existence of the optimal estimator allows a direct comparison of the lower bound, the error covariance of the optimal estimator, and the error covariance of suboptimal estimators for various parameter values and signal-to-noise ratios. Results of Monte-Carlo simulations are presented in Section IV.

II. Optimal and Suboptimal Estimators

The optimal finite dimensional estimator for the system (4)-(6) is constructed as follows [11]-[12]. First, the state x of the linear system (4) is augmented with the state η satisfying

$$\begin{bmatrix} \dot{\eta}_1(t) \\ \dot{\eta}_2(t) \end{bmatrix} = \begin{bmatrix} \alpha - \gamma - P_{11}^{-1}(t) & 0 \\ 0 & \beta - \gamma - P_{22}^{-1}(t) \end{bmatrix} \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \end{bmatrix} + \begin{bmatrix} x_2(t) \\ x_1(t) \end{bmatrix}$$

$$\eta_1(0) = \eta_2(0) = 0 \tag{7}$$

where $P_{ii}(t)$, $i = 1,2$, are the elements of the (nonrandom) Kalman-Bucy filter error covariance matrix $P(t)$ for the linear system (4),(6) ($P(t)$ is obviously diagonal). *The Kalman-Bucy filter for the linear system* (4),(7), with observations (6), computes the conditional expectations $\hat{x}(t|t)$ and $\hat{\eta}(t|t)$. Finally, $\hat{y}(t|t)$ is computed according to

$$\begin{aligned} d\hat{y}(t|t) &= \left[-\gamma\hat{y}(t|t) + \hat{x}_1(t|t)\hat{x}_2(t|t)\right]dt + \frac{1}{r}\hat{\eta}'(t|t)P(t)d\nu(t) \\ \hat{y}(0|0) &= E\left[y(0)\right] \end{aligned} \tag{8}$$

where the innovations process ν is given by

$$d\nu(t) = dz(t) - \hat{x}(t|t)dt, \tag{9}$$

$dz(t) = [dz_1(t), dz_2(t)]'$, and the prime denotes matrix transpose. Hence, the structure of the nonlinearity in (5) allows the computation of $\hat{y}(t|t)$ with a five-state $(\hat{x},\hat{\eta},\hat{y})$ estimator.

The existence of a nonlinear time-invariant steady-state estimator for this problem is also demonstrated in [12]. The assumption which assures this result is the controllability and observability of the original linear system (4),(6), and some additional derivations show the existence of a steady-state estimator for η and y.

The EKF [4] for the state of the nonlinear component of the system (4)-(6) is given by

$$d\tilde{y}(t|t) = \left[-\gamma\tilde{y}(t|t) + \hat{x}_1(t|t)\hat{x}_2(t|t)\right]dt + \frac{1}{r}\left[K_1(t), K_2(t)\right]d\nu(t)$$
$$\tilde{y}(0|0) = E[y(0)] \tag{10}$$

where the "gain" terms satisfy

$$\dot{K}_1(t) = \left[-\alpha - \gamma - \frac{1}{r}P_{11}(t)\right]K_1(t) + P_{11}(t)\hat{x}_2(t|t) \tag{11}$$

$$\dot{K}_2(t) = \left[-\beta - \gamma - \frac{1}{r}P_{22}(t)\right]K_2(t) + P_{22}(t)\hat{x}_1(t|t) \tag{12}$$

$K_1(0) = K_2(0) = 0$, $\hat{x}_1(t|t)$ and $\hat{x}_2(t|t)$ are generated by the Kalman-Bucy filter, and $d\nu(t)$ is given in (9). Notice that the EKF (10) and the optimal estimator (8) differ only in their gain terms. The question of whether constant gains in (10) will suffice is also studied (see Section IV), by utilizing the constant gain extended Kalman filter (CGEKF) of Safonov and Athans [18]; however, their results are not strictly applicable to the system (4)-(6), because polynomial nonlinearities do not satisfy the finite incremental gain conditions of [18].

Now consider the best linear estimator (BLE). Such an estimator can be constructed by first finding a new state equation which is linear, with additive white Gaussian noise, and has the same first two moments (mean and covariance) as those of the original system (4)-(5); then the Kalman-Bucy filter for the new linear state and the observation (6) computes the BLE ([19], p. 152). Let $\bar{x} = [\bar{x}_1, \bar{x}_2, \bar{y}]'$ be the new state. We first compute the mean and covariance propagation equations for the original state $[x',y]'$. Let $m_{x_1}(t) = E[x_1(t)]$, $m_{x_2}(t) = E[x_2(t)]$, and $m_y(t) = E[y(t)]$. Then,

$$\dot{m}_{x_1}(t) = -\alpha m_{x_1}(t) \quad ; \quad m_{x_1}(0) = E[x_1(0)] \tag{13}$$

$$\dot{m}_{x_2}(t) = -\beta m_{x_2}(t) \quad ; \quad m_{x_2}(0) = E[x_2(0)] \tag{14}$$

$$\dot{m}_y(t) = -\gamma m_y(t) + m_{x_1}(t) m_{x_2}(t); \quad m_y(0) = E[y(0)] \tag{15}$$

Let $\Gamma(t)$ be the covariance matrix of $[x',y]'$. (Γ is symmetric and $\Gamma_{x_1x_2} = 0$). Applying Ito's differentiation rule [4] to the elements of $\Gamma(t)$ and taking expectations in the resulting differentials, we have

$$\dot{\Gamma}_{x_1x_1}(t) = q - 2\alpha\Gamma_{x_1x_1}(t) \quad ; \quad \Gamma_{x_1x_1}(0) = \mathrm{var}(x_1(0)) \tag{16}$$

$$\dot{\Gamma}_{x_2x_2}(t) = q - 2\beta\Gamma_{x_2x_2}(t) \quad ; \quad \Gamma_{x_2x_2}(0) = \mathrm{var}(x_2(0)) \tag{17}$$

$$\dot{\Gamma}_{x_1y}(t) = (-\alpha+\gamma)\Gamma_{x_1y}(t) + m_{x_2}(t)\Gamma_{x_1x_1}(t); \quad \Gamma_{x_1y}(0) = 0 \tag{18}$$

$$\dot{\Gamma}_{x_2y}(t) = (-\beta+\gamma)\Gamma_{x_2y}(t) + m_{x_1}(t)\Gamma_{x_2x_2}(t); \quad \Gamma_{x_2y}(0) = 0 \tag{19}$$

$$\begin{aligned} \dot{\Gamma}_{yy}(t) &= -2\gamma\Gamma_{yy}(t) + 2E\left[y(t)x_1(t)x_2(t)\right] - 2m_{x_1}(t)m_{x_2}(t)m_y(t) \\ \Gamma_{yy}(0) &= \mathrm{var}(y(0)) \end{aligned} \tag{20}$$

The expectation of $x_1(t)x_2(t)y(t)$ in (20) is computed as follows.

$$\begin{aligned} E\left[y(t)x_1(t)x_2(t)\right] &= E\left\{\left[e^{-\gamma t}\, y(0) + \int_0^t e^{-\gamma(t-\tau)}\, x_1(\tau)x_2(\tau)d\tau\right]x_1(t)x_2(t)\right\} \\ &= e^{-\gamma t}\, m_y(0)m_{x_1}(t)m_{x_2}(t) + \int_0^t e^{-\gamma(t-\tau)}\, E\left[x_1(t)x_1(\tau)\right] \\ &\qquad \cdot E\left[x_2(t)x_2(\tau)\right]d\tau \\ &\triangleq e^{-\gamma t}\, m_y(0)m_{x_1}(t)m_{x_2}(t) + Q(t) \end{aligned} \tag{21}$$

where Q(t) can be expressed in differential form

$$\dot{Q}(t) = -(\alpha+\beta+\gamma)Q(t) + \left[m_{x_1}^2(t) + \Gamma_{x_1x_1}(t)\right]\left[m_{x_2}^2(t) + \Gamma_{x_2x_2}(t)\right]$$

$$Q(0) = 0 \tag{22}$$

Substituting (21) into (20), we have, if $\gamma \neq \alpha+\beta$,

$$\dot{\Gamma}_{yy}(t) = -2\gamma\Gamma_{yy}(t) + 2\left\{Q(t) - \frac{m_{x_1}^2(t)m_{x_2}^2(t)}{\gamma-\alpha-\beta}\left[1 - e^{-(\gamma-\alpha-\beta)t}\right]\right\}$$

$$\triangleq -2\gamma\Gamma_{yy}(t) + 2\bar{Q}(t) \tag{23}$$

and $\bar{Q}(t) = Q(t) - m_{x_1}^2(0)m_{x_2}^2(0)t\ e^{-(\gamma+\alpha+\beta)t}$, if $\gamma = \alpha+\beta$. The linear state equation for $\bar{x}$ is thus formulated to be

$$d\bar{x}(t) = \bar{F}\bar{x}(t)dt + \bar{G}(t)d\bar{w}(t) + \left[0,0,m_{x_1}(t)m_{x_2}(t)\right]' \tag{24}$$

where

$$\bar{F} = \begin{bmatrix} -\alpha & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & -\gamma \end{bmatrix}, \quad \bar{G}(t) = \begin{bmatrix} q^{1/2} & 0 & 0 \\ 0 & q^{1/2} & 0 \\ 0 & 0 & \sqrt{2\bar{Q}(t)} \end{bmatrix},$$

and $\bar{w}(t)$ is a Brownian motion of dimension 3 with $\tilde{Q}(t)dt = E\{d\bar{w}(t)d\bar{w}'(t)\}$. Moreover, $\bar{G}(t)\tilde{Q}(t)\bar{G}'(t)$ should satisfy

$$\bar{G}(t)\tilde{Q}(t)\bar{G}'(t) = \begin{bmatrix} q & 0 & m_{x_2}(t)\Gamma_{x_1x_1}(t) \\ 0 & q & m_{x_1}(t)\Gamma_{x_2x_2}(t) \\ m_{x_2}(t)\Gamma_{x_1x_1}(t) & m_{x_1}(t)\Gamma_{x_2x_2}(t) & 2\bar{Q}(t) \end{bmatrix} \tag{25}$$

Then the mean and covariance propagation equations for $\bar{x}(t)$ have the same forms as those given in equations (13)-(19),(23). Observe that in (24), $\bar{x}_1(t) = x_1(t)$ and $\bar{x}_2(t) = x_2(t)$; only the nonlinear component of (5) has been altered. The BLE is the Kalman-Bucy filter for the system (24) with observation (6).

III. Performance of Estimators and Estimation Lower Bound

In this section we shall derive the error covariance propagation equations associated with the optimal and suboptimal estimators. We are interested in computing the error covariance of y, since the error covariance of x is readily implemented via the Kalman-Bucy filter. Though many lower bounds have been proposed to facilitate the evaluation of suboptimal estimators, for example, Snyder and Rhodes [5],[6] for Gaussian state processes, Gilman and Rhodes [7],[8] for systems with cone-bounded nonlinearities, and Zakai and his colleagues [9],[10] for fairly general problems in both discrete and continuous time, only the lower bound derived by Bobrovsky and Zakai [10] is applicable to the system (4)-(6). We shall investigate the tightness of Bobrovsky and Zakai's lower bound on estimator performance -- that is, how close is the lower bound to the covariance of the <u>optimal</u> estimator?

The error covariance of y is the expected value of the conditional error covariance $E[(y(t)-\hat{y}(t|t))^2|z^t]$. Taking the expectations in the equation ([4], equation (6.100)) satisfied by the conditional error covariance, we obtain the error covariance propagation equation associated with $\hat{y}(t|t)$:

$$\begin{aligned}\dot{\Sigma}(t) &= -2\gamma\Sigma(t) + 2E\left[E^t[y(t)x_1(t)x_2(t)] - E^t[y(t)]E^t[x_1(t)x_2(t)]\right]\\ &\quad - \frac{1}{r}\,E\left[\Sigma_1^2(t) + \Sigma_2^2(t)\right]\\ \Sigma(0) &= \mathrm{var}(y(0)) \end{aligned} \qquad (26)$$

where $\Sigma_i(t) = E[(x_i(t)-\hat{x}_i(t|t))(y(t)-\hat{y}(t|t))|z^t]$, $i=1,2$. In the sequel, we assume that $x_1(0)$ and $x_2(0)$ are independent, zero-mean, Gaussian random variables. It is easy to show that $\Sigma_1(t) = \Sigma_2(t) = 0$, for $x_1(t)$ and $x_2(t)$ are independent, zero-mean, Gaussian processes. The other expectation in (26) is computed as follows:

$$E^t\Big[y(t)x_1(t)x_2(t)\Big] - E^t\Big[y(t)\Big]\; E^t\Big[x_1(t)x_2(t)\Big]$$

$$= \int_0^t e^{-\gamma(t-s)} \Big\{ E^t\Big[x_1(s)x_2(s)x_1(t)x_2(t)\Big] - E^t\Big[x_1(s)x_2(s)\Big]\; E^t\Big[x_1(t)x_2(t)\Big] \Big\} ds$$

$$= \int_0^t e^{-\gamma(t-s)} \Big[P_{11}(s,t,t)\hat{x}_2(s|t)\hat{x}_2(t|t) + P_{22}(s,t,t)\hat{x}_1(s|t)\hat{x}_1(t|t) + P_{11}(s,t,t)P_{22}(s,t,t)\Big]\, ds$$

$$= P_{11}(t)\hat{\eta}_1(t|t)\hat{x}_2(t|t) + P_{22}(t)\hat{\eta}_2(t|t)\hat{x}_1(t|t) + \int_0^t e^{-\gamma(t-s)}\, P_{11}(s,t,t)P_{22}(s,t,t)ds \tag{27}$$

where $P_{ij}(s,t,t) = E[(x_i(s) - \hat{x}_i(s|t))(x_j(t) - \hat{x}_j(t|t))|z^t]$, $i,j = 1,2$, are the nonrandom conditional cross-covariances defined in [12, Lemma 2.1]; here, $P_{12}(s,t,t) = P_{21}(s,t,t) = 0$. A crucial component in computing (27) is the use of [12, Lemma B.1], which expresses the higher order moment $E^t[x_1(s)x_2(s)x_1(t)x_2(t)]$ of a Gaussian distribution in terms of lower order moments. We need only compute the expected values of $\hat{\eta}_1(t|t)\hat{x}_2(t|t)$ and $\hat{\eta}_2(t|t)\hat{x}_1(t|t)$, since the last integral in (27) is nonrandom.

The state equation of $\hat{\xi}(t|t) \triangleq [\hat{x}'(t|t),\hat{\eta}'(t|t)]'$ is the Kalman-Bucy filter for the augmented system (4),(7) with observation (6). $\hat{\xi}$ satisfies

$$d\hat{\xi}(t|t) = C(t)\; \hat{\xi}(t|t) + \frac{1}{r}\; D(t)d\nu(t) \tag{28}$$

where

$$C(t) \triangleq \begin{bmatrix} -\alpha & 0 & 0 & 0 \\ 0 & -\beta & 0 & 0 \\ 0 & 1 & \alpha-\gamma-qP_{11}^{-1}(t) & 0 \\ 1 & 0 & 0 & \beta-\gamma-qP_{22}^{-1}(t) \end{bmatrix}, \quad D(t) \triangleq \begin{bmatrix} S_{11}(t) & 0 \\ 0 & S_{22}(t) \\ 0 & S_{23}(t) \\ S_{14}(t) & 0 \end{bmatrix},$$

and $S(t)$ is the conditional error covariance matrix of $\xi(t)$ given z^t; it satisfies the Riccati equation, and $S_{11}(t) = P_{11}(t)$ and $S_{22}(t) = P_{22}(t)$. Since the innovations process ν is a standard Brownian motion process with unit variance [2], the covariance $\Phi(t)$ of $\hat{\xi}$ is easily computed according to

$$\dot{\Phi}(t) = C(t)\Phi(t) + \Phi(t)C'(t) + \frac{1}{r} D(t)D'(t) . \tag{29}$$

$E[\hat{\eta}_1(t|t)\hat{x}_2(t|t)]$ and $E[\hat{\eta}_2(t|t)\hat{x}_1(t|t)]$ are, respectively, the (2,3) and (1,4) elements of $\Phi(t)$. Expressing the last integral in (27) in differential form, we have

$$\int_0^t e^{-\gamma(t-s)} P_{11}(s,t,t)P_{22}(s,t,t)ds = P_{11}(t)P_{22}(t)\delta(t)$$

where $\delta(t)$ satisfies

$$\dot{\delta}(t) = 1 + \left[\alpha+\beta-\gamma-qP_{11}^{-1}(t) - qP_{22}^{-1}(t)\right]\delta(t); \qquad \delta(0) = 0 \tag{30}$$

Finally, equation (26) is rewritten as

$$\begin{aligned} \dot{\Sigma}(t) &= -2\gamma\Sigma(t) + 2\left\{P_{11}(t)\Phi_{23}(t) + P_{22}(t)\Phi_{14}(t) + P_{11}(t)P_{22}(t)\delta(t)\right\} \\ &\triangleq -2\gamma\Sigma(t) + 2\theta(t) \end{aligned} \tag{31}$$

where $\theta(t) \geq 0$ for all $t \geq 0$, is a function depending on the system parameters (α,β,γ) and noise covariances (q and r). As shown in [12], the augmented linear system (4),(7) is time-invariant in steady state and $S(t)$ has a unique positive-definite steady-state solution S. Similarly, equations (29),(30) have steady-state solutions Φ and δ, respectively, where Φ is positive definite and δ is positive. Hence, at steady state, the solution of (31) is $\Sigma = \theta/\gamma$.

The lower bound [10] for the system (4)-(6) is obtained by applying the Kalman-Bucy filter to the following linearized system

$$dh(t) = A(t)h(t)dt + G(t)dw(t) \tag{33}$$

$$dk(t) = B(t)h(t)dt + r^{\frac{1}{2}}\,dv(t) \tag{34}$$

where $h(t)$ and $k(t)$ are 3- and 2-dimensional vectors, respectively, and $w(t)$ and $v(t)$ are independent standard Brownian motions of dimensions 3 and 2, respectively. $A(t)$, $B(t)$, $G(t)$ are 3×3, 2×3, 3×3 matrices, respectively:

$$A(t) = \begin{bmatrix} -\alpha & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & -\gamma \end{bmatrix}, \quad G(t) = \begin{bmatrix} q & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and $B(t)$ is the solution of

$$\frac{1}{r}B'(t)B(t) = \begin{bmatrix} \frac{1}{r} + \frac{1}{q}E\left[x_2^2(t)\right] & 0 & 0 \\ 0 & \frac{1}{r} + \frac{1}{q}E\left[x_1^2(t)\right] & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$h(0)$ is assumed to be a zero-mean, Gaussian random vector with nonsingular variance, and $k(0) = 0$.

The lower bound $\Sigma_{LB}(t)$ of Bobrovsky and Zakai, which lower bounds the error covariance of any estimator of y, is the (3,3) element of the error covariance matrix for the system (33)-(34). $\Sigma_{LB}(t)$ satisfies

$$\dot{\Sigma}_{LB}(t) = -2\gamma\Sigma_{LB}(t) \tag{35}$$

Comparing (31) and (35), we have $\Sigma_{LB}(t) \leq \Sigma(t)$ for all t, as expected. The bound does not depend on the measurement noise covariance r; for some values of r it is far from the optimal error covariance (see Section IV), since $\theta(t)$ in (31) is nonzero and can in fact be large. This demonstrates that the lower bound of Bobrovsky and Zakai can be quite loose in some cases. This was to be expected, since the bound is based upon the Cramer-Rao lower bound, which can be loose if the signal-to-noise ratio is small [20]. However, in steady-state $\Sigma(t) \to \Sigma_{LB}(t)$ as $r \to 0$, which can be shown as follows. Based on the discussions following (31), the steady state solution of $\Sigma(t)$ is

$$\Sigma = \frac{\theta}{\gamma} = \frac{1}{\gamma}\left\{P_{11}\Phi_{23} + P_{22}\Phi_{14} + P_{11}P_{22}\,\delta\right\}$$

where

$$P_{11} = -\alpha r + \sqrt{\alpha^2 r^2 + qr} \quad , \quad P_{22} = -\beta r + \sqrt{\beta^2 r^2 + qr} \quad ,$$

$$S_{14} = \frac{P_{11}P_{22}}{(P_{11}P_{22}/r) + (\alpha-\beta+\gamma)P_{22} + q} \, , \quad S_{23} = \frac{P_{11}P_{22}}{(P_{11}P_{22}/r) + (-\alpha+\beta+\gamma)P_{11} + q} \, ,$$

$$\Phi_{23} = \frac{(P_{11}P_{22}^2/r) + 2\beta(P_{11}P_{22}/r)S_{23}}{2\beta[(-\alpha+\beta+\gamma)P_{11} + q]} \, , \quad \Phi_{14} = \frac{(P_{11}^2P_{22}/r) + 2\alpha(P_{11}P_{22}/r)S_{14}}{2\alpha[(\alpha-\beta+\gamma)P_{22} + q]} \, ,$$

and

$$\delta = \frac{1}{-\alpha-\beta+\gamma+q(1/P_{11} + 1/P_{22})} \, .$$

We can easily see that as $r \to 0$, $\Sigma \to \Sigma_{LB}(=0)$; thus, at least at steady state, this result confirms the prediction by Bobrovsky and Zakai [10].

The EKF generates an "approximation" $\tilde{\Sigma}(t)$ to the conditional error covariance $E[(y(t) - \tilde{y}(t|t))^2 | Z^t]$, which is given by

$$\dot{\tilde{\Sigma}}(t) = -2\gamma\tilde{\Sigma}(t) + 2K_1(t)\hat{x}_2(t|t) + 2K_2(t)\hat{x}_1(t|t) - [K_1(t)]^2 - [K_2(t)]^2 \quad (36)$$

Taking expectations in (36), we find that the EKF approximation $E[\tilde{\Sigma}(t)]$ to the error covariance $E[(y(t) - \tilde{y}(t|t)^2]$ satisfies

$$\frac{d}{dt} E[\tilde{\Sigma}(t)] = -2\gamma E[\tilde{\Sigma}(t)] \quad (37)$$

Thus $\Sigma_{LB}(t) = E[\tilde{\Sigma}(t)] \leq \Sigma(t)$, indicating that $E[\tilde{\Sigma}(t)]$ may not be a very good approximation to the error covariance $E[(y(t) - \tilde{y}(t|t))^2]$ of the EKF; for, since $\hat{y}(t|t)$ of (8) is the optimal mean-square error estimate, it is clear that in fact $\Sigma(t) \leq E[(y(t) - \tilde{y}(t|t))^2]$. Hence, as is well known, one should exercise caution in using $E[\tilde{\Sigma}(t)]$ as an indicator of EKF performance.

Denoting $\bar{\Sigma}(t) \triangleq E[(y(t) - \bar{y}(t|t))^2]$, the error covariance matrix associated with $\bar{y}(t|t)$ of the BLE, we have

$$\dot{\bar{\Sigma}}(t) = -2\gamma\bar{\Sigma}(t) + 2\bar{Q}(t) - \frac{\bar{\Sigma}_1(t) + \bar{\Sigma}_2(t)}{r}; \quad \bar{\Sigma}(0) = \mathrm{var}(y(0)) \quad (38)$$

where

$$\dot{\bar{\Sigma}}_1(t) = -(\alpha+\gamma)\bar{\Sigma}_1(t) + m_{x_2}(t)\Gamma_{x_1x_1}(t) - \frac{P_{11}(t)\bar{\Sigma}_1(t)}{r}; \quad \bar{\Sigma}_1(0) = 0 \quad (39)$$

$$\dot{\bar{\Sigma}}_2(t) = -(\beta+\gamma)\bar{\Sigma}_2(t) + m_{x_1}(t)\Gamma_{x_2x_2}(t) - \frac{P_{22}(t)\bar{\Sigma}_2(t)}{r}; \quad \bar{\Sigma}_2(0) = 0 \quad (40)$$

and $\bar{\Sigma}_i(t) = E[(x_i(t) - \bar{x}_i(t|t))(y(t) - \bar{y}(t|t))]$, $i = 1,2$.

Consider now the stability of (38)-(40). First, as $t\to\infty$, $m_{x_1}(t)$ and $m_{x_2}(t)$ approach zero. Also, since (4),(6) is controllable and observable, $P_{11}(t)$ and $P_{22}(t)$ approach positive constant values. Thus [24, p. 113], $\bar{\Sigma}_1(t)$ and $\bar{\Sigma}_2(t)$ approach zero as $t\to\infty$. Comparing (38) and (23), we find that $\bar{\Sigma}(t)$ has the same performance as that of $\Gamma_{yy}(t)$ (the a priori covariance) as $t\to\infty$. Furthermore, if $x_1(0)$ and $x_2(0)$ are zero-mean, $m_{x_1}(t)$, $m_{x_2}(t)$, $\bar{\Sigma}_1(t)$, and $\bar{\Sigma}_2(t)$ will be zero for all $t\geq 0$. Thus, the observation process $z(t)$ is uncorrelated with $y(t)$ in this case, and $\bar{y}(t|t)$ is just the a priori mean $m_y(t)$; that is, the BLE for $y(t)$ does not take advantage of the observations during the course of estimation. The error covariance $\bar{\Sigma}(t)$ of the BLE is also the a priori covariance $\Gamma_{yy}(t)$ which, at steady state, is

$$\Gamma_{yy} = \bar{\Sigma} = \frac{q^2}{4\alpha\beta\gamma(\alpha+\beta+\gamma)} \tag{41}$$

Although in this case $y(t)$ is uncorrelated with $z(t)$, it is not independent of $z(t)$; hence the optimal estimator which computes the conditional mean $\hat{y}(t|t)$ can in fact perform quite well, as is demonstrated in Section IV. It should also be pointed out that for the system (4)-(6) the optimal estimator is also the best quadratic estimator; that is, it is the best estimator whose input/output map can be expressed as a Volterra series of order 2 with input as the innovations process $\nu(t)$ (see [21]).

In [10] Bobrovsky and Zakai showed that for a scalar problem and for sufficiently small values of r, optimal linear filtering yielded practically the same filtering error as that of the optimal nonlinear filtering. This is not the case for our multivariable problem. In fact, for sufficiently large q and small r, $\bar{\Sigma}$ is much greater than Σ and Σ_{LB} (see Section IV).

IV. Simulation Results

In order to compare and evaluate the performance of the optimal estimator, the EKF, the CGEKF, the BLE, and the lower bound, digital Monte Carlo simulations were employed. In this section it is assumed that $x_1(0)$ and $x(0)$ have zero mean. Since the error covariance propagation equations for the optimal estimator (31), the BLE (38), and the lower bound (35) are ordinary differential equations, they were computed off-line and stored. Simulations were conducted on the EKF (10), the CGEKF, and the optimal estimator (8). Identical noise sequences were used to allow direct comparison. Our approach to the statistical analysis of the Monte Carlo simulations parallels that of Bucy and his associates [22]. The mean-square error

$$\mu = E[(y(t) - y^*(t))^2] \tag{42}$$

where $y^*(t)$ denotes the estimate (e.g., $\hat{y}(t|t)$ or $\tilde{y}(t|t)$), was used as the performance measure in the simulation. Suppose that $\{y_n^*\}$ and $\{y_n\}$, $n = 1,\ldots,N$, are sequences of independent realizations of $y_n^*(t)$ and $y(t)$, respectively. Then the statistic

$$\hat{\mu} = \frac{1}{N} \sum_{n=1}^{N} (y_n - y_n^*)^2 \tag{43}$$

is an approximation to μ for sufficiently large N.

In the experiments, the parameters were varied as follows:

$\alpha, \beta, \gamma = 0.9$

$q = 2$

$r = 0.04$, 0.2 and 1

75 sample paths, each of which contained 5000 steps of length 0.001s, were run in each simulation. The constant gains in the CGEKF were chosen to be the average gain values (the gains in (10)) over 75 sample paths; both gain values were quite small (we also simulated the zero-gain EKF; the simulation result was almost the same as that for nonzero gains). Figs. 1-3 display, for three cases, graphs of mean-square error (averaged over 75 sample paths) of the optimal estimator, the suboptimal estimators (EKF, CGEKF, BLE), and the lower bound.

As expected, the steady state optimal error covariance approaches the lower bound for very small values of r (=0.04) and is greater than the lower bound for large r (=1). The BLE remains unchanged in the three figures, since it is just the a priori mean. The EKF performance is quite close to the simulated optimal error covariance in every simulation run; further simulations [23] have indicated that the gain terms (see (8) and (10)) of the optimal estimator and the EKF are almost equal. The performance of the CGEKF is somewhat less effective than that of the EKF, but is far superior to the BLE performance. The estimate $\tilde{\bar{y}}(t|t)$ of the CGEKF is the result of passing the <u>a posteriori</u> means $\hat{x}_1(t|t)$ and $\hat{x}_2(t|t)$ through the nonlinear filter (10) with gains set to near-zero values, while the $\bar{y}(t|t)$ of the BLE results from passing the <u>a priori</u> means $m_{x_1}(t)$ and $m_{x_2}(t)$ through (15). This of course explains why the CGEKF, which utilizes the observations, has better performance than the BLE. It should be noted that the difference between the averaged mean-square error of the optimal estimator and the actual optimal error covariance is due to the fact that averaging over 75 sample paths does not give a good approximation to the expectation.

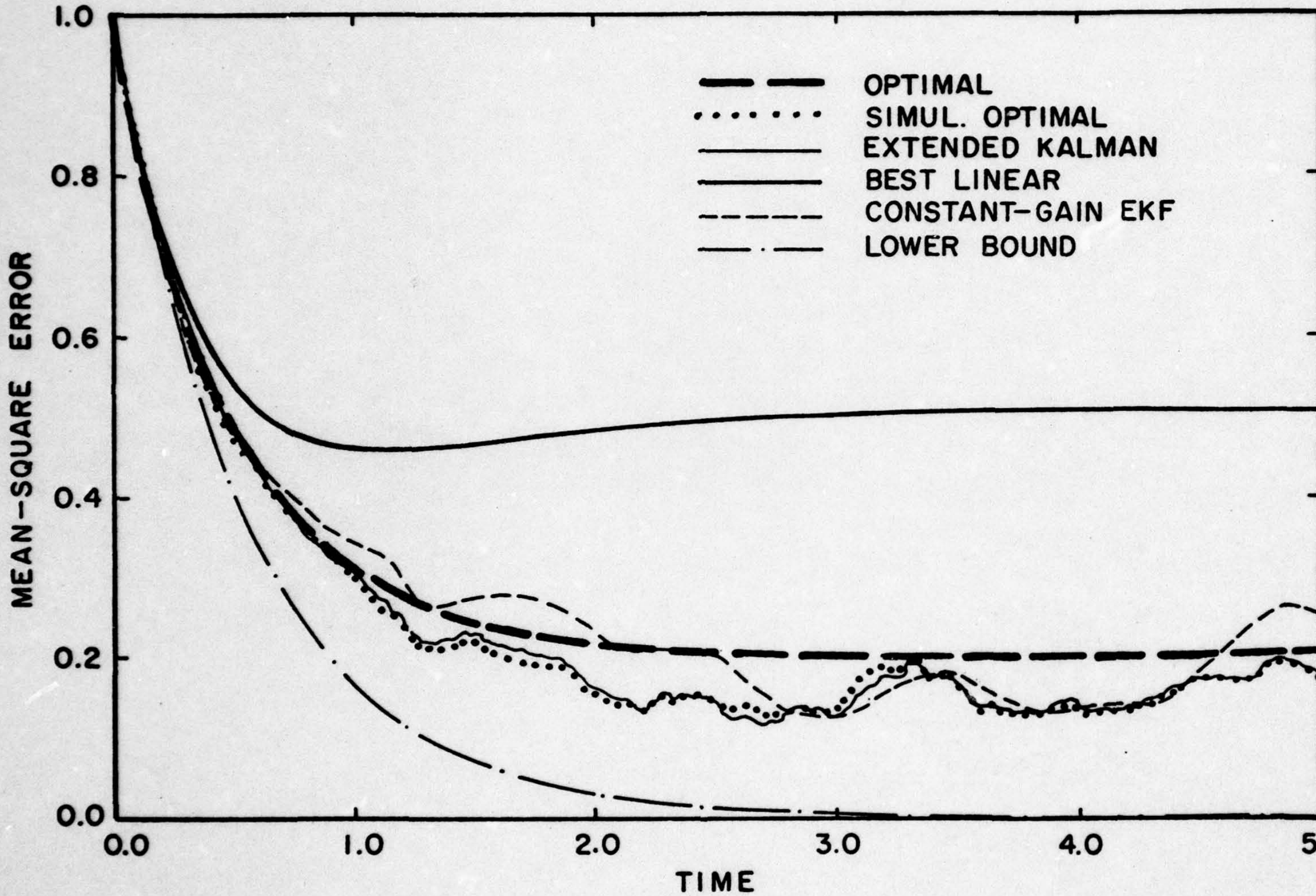


Figure 1. Mean-square error: $q = 2.0$, $r = 0.2$

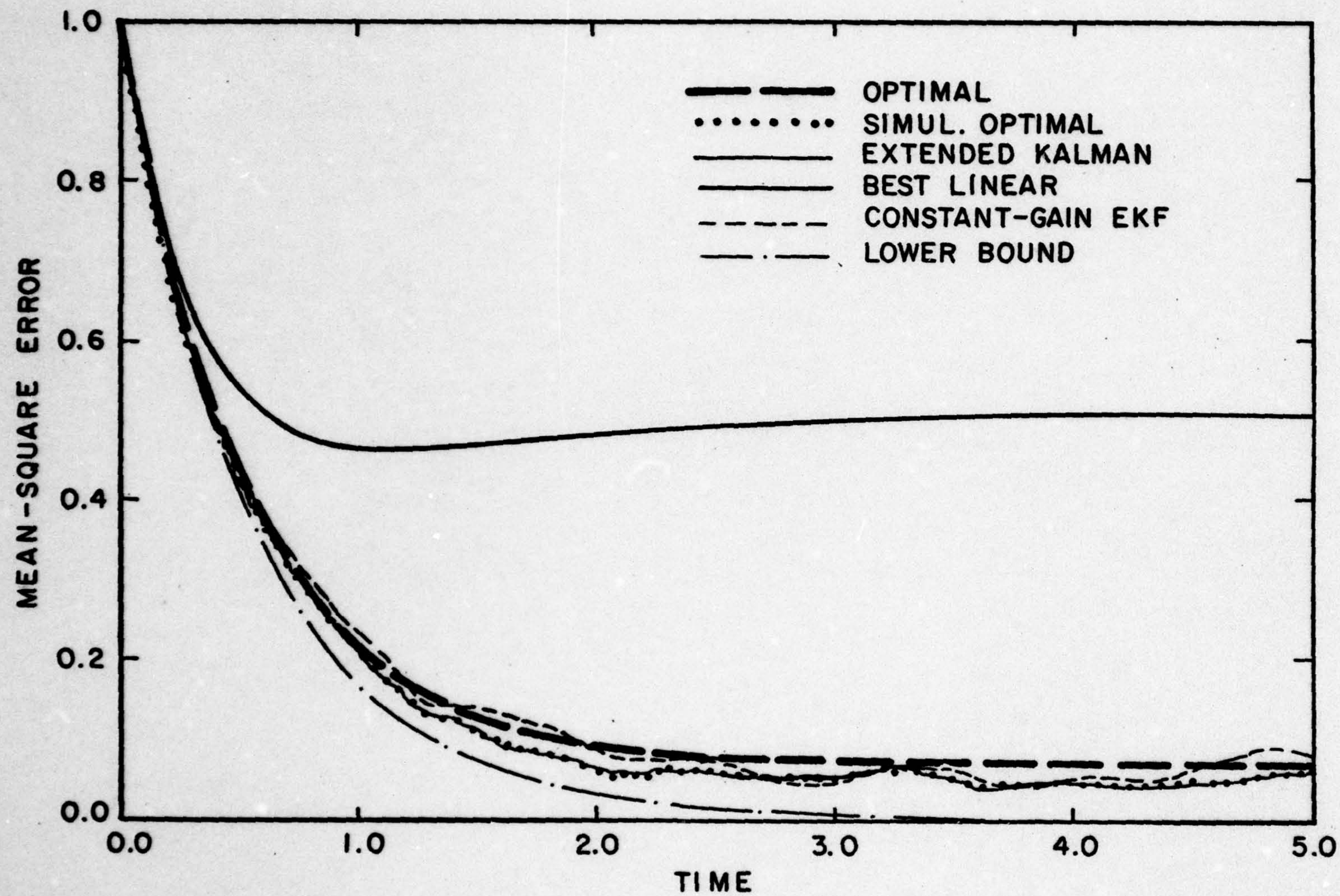


Figure 2. Mean-square error: $q = 2.0$, $r = 0.04$

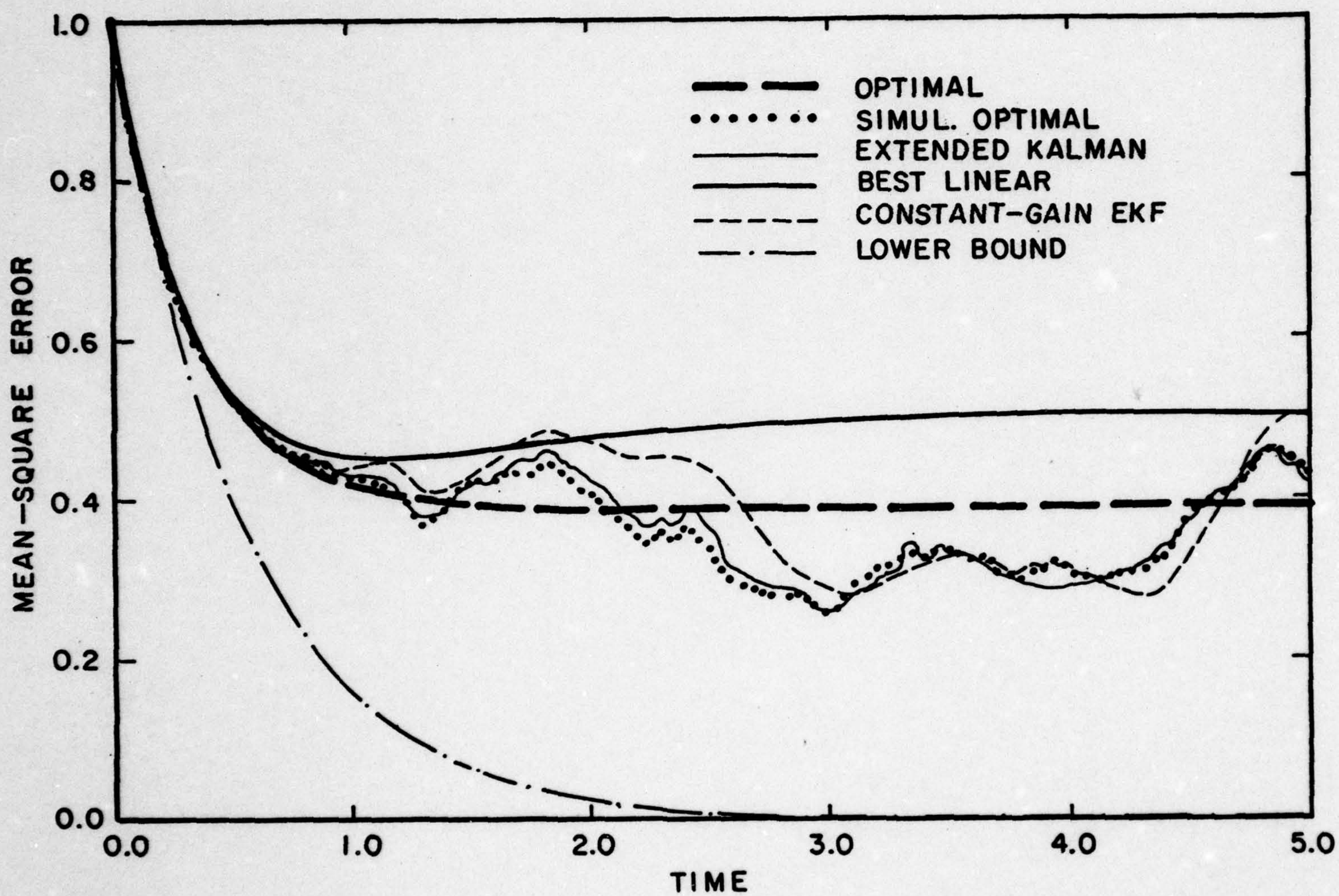


Figure 3. Mean-square error: q = 2.0, r = 1.0

V. Concluding Remarks

A nonlinear system with a finite dimensional optimal nonlinear estimator has been considered. Based on such a system, optimal and suboptimal estimators and an estimation lower bound have been studied. The optimal estimator is used as a criterion in evaluating the performance of the suboptimal estimators and the lower bound. Simulation results indicate that the performance of the EKF is as good as that of the optimal estimator, and the Bobrovsky and Zakai lower bound is tight for very high signal-to-noise ratio ($r \to 0$) but is less effective for large values of state and observation noises. As far as suboptimal filter design is concerned, the CGEKF is probably preferable, in most of the cases studied, to the optimal estimator and the EKF, due to its simple computational requirements. In fact, a naive estimator which passes $\hat{x}_1(t|t)$ and $\hat{x}_2(t|t)$ through a model of the nonlinear system (5) performs almost as well.

References

1. H. J. Kushner, "Dynamical equations for optimal nonlinear filtering," *J. Differential Equations*, vol. 3, pp. 179-190, 1967.

2. M. Fujisaki, G. Kallianpur, and H. Kunita, "Stochastic differential equations for the nonlinear filtering problem," *Osaka J. Math.*, vol. 9, pp. 19-40, 1972.

3. R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *J. Basic Engr. (Tran. ASME)*, vol. 83D, pp. 95-108, 1961.

4. A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.

5. D. L. Snyder and I. B. Rhodes, "Filtering and control performance bounds with implications on asymptotic separation," *Automatica*, vol. 8, pp. 747-753, 1972.

6. I. B. Rhodes and D. L. Snyder, "Estimation and control performance for space-time point-process observations," *IEEE Transactions on Automatic Control*, vol. AC-22, pp. 338-346, June 1977.

7. A. S. Gilman and I. B. Rhodes, "Cone-bounded nonlinearities and mean-square bounds," *IEEE Transactions on Automatic Control*, vol. AC-18, pp. 260-265, June 1973.

8. I. B. Rhodes and A. S. Gilman, "Cone-bounded nonlinearities and mean-square bounds: Estimation lower bounds," *IEEE Transactions on Automatic Control*, vol. AC-20, pp. 632-642, October 1975.

9. M. Zakai and J. Ziv, "Lower and upper bounds on the optimal filtering error of certain diffusion process," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 325-331, May 1972.

10. B. Z. Bobrovsky and M. Zakai, "A lower bound on the estimation error for certain diffusion process," *IEEE Transactions on Information Theory*, vol. IT-22, pp. 45-52, January 1976.

11. S. I. Marcus, "Estimation and analysis of nonlinear stochastic systems," Ph.D. dissertation, Dept. of Electrical Engineering, M.I.T., Cambridge, MA, June 1975.

12. S. I. Marcus and A. S. Willsky, "Algebraic structure and finite dimensional nonlinear estimation," *SIAM J. Math. Analysis*, vol. 9, pp. 312-327, April 1978.

13. A. S. Willsky and S. I. Marcus, "Estimation of bilinear stochastic systems," *Variable Structure Systems with Applications to Economics and Biology,* R. R. Mohler and A. Ruberti, eds., New York: Springer-Verlag, pp. 116-137, 1975.

14. J. T. Lo and A. S. Willsky, "Estimation for rotational processes with one degree of freedom I: Introduction and continuous time processes," *IEEE Transactions on Automatic Control,* vol. AC-20, pp. 10-21, 1975.

15. A. S. Willsky, "Some estimation problems on Lie groups," *Geometric Methods in System Theory,* R. W. Brockett and D. Q. Mayne, eds., Reidel, The Netherlands, 1973.

16. R. W. Brockett, "Volterra series and geometric control theory," *Automatica,* vol. 12, pp. 167-176, 1976.

17. R. W. Brockett, "Nonlinear systems and differential geometry," *Proceedings of the IEEE,* vol. 64, pp. 61-72, January 1976.

18. M. G. Safanov and M. Athans, "Robustness and computational aspects of nonlinear stochastic estimators and regulators," *IEEE Transactions on Automatic Control,* vol. AC-23, pp. 717-725, August 1978.

19. R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes II.* New York: Springer-Verlag, 1978.

20. D. J. Sakrison, *Notes on Analog Communication.* New York: Van Nostrand Reinhold, 1970.

21. S. I. Marcus, S. K. Mitter, and D. Ocone, "Finite dimensional nonlinear estimation in continuous and discrete time," presented at the International Conference on Analysis and Optimization of Stochastic Systems, Oxford, September 6-8, 1978.

22. R. S. Bucy, C. Hecht, and K. O. Senne, "An engineer's guide to building nonlinear filters," F. J. Seiler Research Lab., Project 7904-00-37, May 1972.

23. C.-H. Liu, "A comparison of optimal and suboptimal estimators and estimation lower bounds," M.S. Thesis, Dept. of Electrical Engineering, University of Texas at Austin, August 1978.

24. J. L. Willems, *Stability Theory of Dynamical Systems.* New York: Wiley, 1970.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) (14) AFOSR-TR-79-0454	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (6) Estimator Performance for a Class of Nonlinear Estimation Problems		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER (15) F49620-77-C-0101
7. AUTHOR(s) (10) Chang-Huan Liu [redacted] Steven I. Marcus		8. CONTRACT OR GRANT NUMBER(s) AFOSR-79-0025 new
9. PERFORMING ORGANIZATION NAME AND ADDRESS The University of Texas at Austin Department of Electrical Engineering Austin, Texas 78712		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F (17) (16) 2304 A1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, D.C. 20332	(11) 23	12. REPORT DATE Mar[redacted]79
		13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 24p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Estimation, Estimation Lower Bounds.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The state estimation problem for a certain class of nonlinear stochastic systems with white Gaussian plant and observation noise is considered. The optimal (minimum variance) estimators for these systems are recursive and finite dimensional. A particular nonlinear system which contains a polynomial nonlinearity is presented. Both optimal and suboptimal estimators and an estimation lower bound for such a system are derived. The performance of the optimal and suboptimal estimators and the lower bound are compared both analytically and by computer simulation.

DD FORM 1 JAN 73 1473

401 997

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)